DOCUMENT RESUME

ED 264 516 CS 008 081

AUTHOR	McConkie, George W.; And Others
TITLE	Estimating Frequency and Size of Effects Due to Experimental Manipulations in Eye Movement Research. Technical Report No. 338.
INSTITUTION	Bolt, Beranek and Newman, Inc., Cambridge, Mass.; Illinois Univ., Urbana. Center for the Study of Reading.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.; National Inst. of Mental Health (DHHS), Rockville, MD.
PUB DATE	Jun 85
CONTRACT	400-81-0030
GRANT	NIMH-MH-32884
NOTE	23p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Eye Fixations; Eye Movements; Higher Education; *Reading Research; Research Problems; Test Validity
IDENTIFIERS	*Frequency Analysis

ABSTRACT

A study of eye movements was conducted to determine whether the visual regions are perceived in their entirety on each fixation or whether the actual region perceived varies from fixation to fixation. The resulting data were used in a frequency of effects analysis. The frequency of effects problem arises once it has been established that an experimental manipulation has produced an effect, and consists of trying to determine the frequency and size of the effect. Text was displayed on a CRT and refreshed every three milliseconds (msec). The subjects' eyes were monitored during reading, with sampling of eye position every msec. During selected fixations, the letters at certain locations were replaced by other letters, thus resulting in erroneous letters being present at specific retinal locations on those fixations. The eye movement data were then analyzed to determine whether errors produced an effect, and if they did, to estimate the frequency with which this occurred. Subjects were 12 college students, each of whom read 16 short passages, which yielded over 1,000 fixations in the control and two experimental conditions. The results indicated that in one instance, a manipulation that produced a 21-msec increase in fixation duration was actually producing a 151-msec increase in only 21% of the instances and was having no effect in the remaining 79% of the cases. The findings suggest that studies of eye movements in reading should use the frequency of effects analysis procedure to determine more accurately the effects of a given manipulation. (HTH)

ED264516

CENTER FOR THE STUDY OF READING

Technical Report No. 338

ESTIMATING FREQUENCY AND SIZE OF EFFECTS DUE TO EXPERIMENTAL MANIPULATIONS IN EYE MOVEMENT RESEARCH

George W. McConkie, David Zola
and Gary S. Wolverton

University of Illinois at Urbana-Champaign

June 1985

University of Illinois
at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

Bolt Beranek and Newman Inc.
10 Moulton Street
Cambridge, Massachusetts 02238

The work upon which this publication is based was performed pursuant to NIMH grant MH 32884 and Contract No. 400-81-0030 of the National Institute of Education. It does not, however, necessarily reflect the views of these agencies. To appear in Groner, McConkie & Menz, Eye Movements and Human Information Processing. Amsterdam, North-Holland.

Abstract

In studies of eye movement behavior, once it has been demonstrated that an experimental manipulation has produced a reliable effect, it is often useful to try to estimate the frequency with which the effect occurred. This paper describes the Frequency of Effects Analysis and illustrates its use with data from a study on characteristics of the perceptual span of adult readers. The results of the analysis indicated that, in one instance, a manipulation which produced a 21 msec increase in fixation duration was actually producing a 151 msec increase in only 21% of the instances, and was having no effect in the remaining 79% of the cases.

Estimating Frequency and Size of Effects

Due to Experimental Manipulations in Eye Movement Research

This paper describes a problem which we have encountered several times recently in our studies of the eye movements of readers, and proposes a direction to take in its solution. We suspect that it is a problem that occurs quite frequently in eye movement research, as well as experimental research in other areas. It is referred to as the frequency of effects problem.

When conducting an experimental study, researchers typically make some stimulus manipulation in one condition, and compare the results with another condition in which the manipulation was not made. In each condition, data are obtained on at least one dependent variable for each trial, which can be represented as frequency distributions of the scores for the two conditions. Statistical techniques are used to determine the likelihood that these two distributions represent samples from two different populations, indicating whether the experimental manipulation produced a difference in the dependent variable. The most frequently used techniques involve calculating the statistical significance of the difference between the means of the two sets of scores. In some cases the difference in the variance of the two distributions is tested, or, through the use of something like the Kolmogorov-Smirnov test, the shapes of the distributions themselves are compared. A variety of other non-parametric tests can also be used. All these are aimed at

determining whether the experimental manipulation produced an effect on the dependent variable.

The Frequency of Effects problem arises when in order to answer some theoretical question, it is necessary to know more than just whether there was an effect, or the average size of that effect. For example, suppose that a study were conducted from which it was determined that an experimental manipulation produced a 20 msec increase in the duration of the next fixation. That is, the frequency distributions of these particular fixations had a mean that was 20 msec greater in the experimental condition than in the control condition. It is tempting to conclude that the response to that manipulation was to increase fixation duration by 20 msec, perhaps as a result of extra processing. Such a conclusion is based on the assumption that the experimental manipulation had an effect on 100% of its occurrences, and that the size of that effect, while perhaps variable, averaged 20 msec. In fact, this assumption has not been tested, and may actually be false. It is equally possible that the manipulation produced an effect on the dependent variable on only 25% of its occurrences, and on those instances the fixation duration was increased by an average of 80 msec. At the same time, there was no effect at all on 75% of the instances. These two possibilities and many others are equally harmonious with the original finding of a 20 msec average effect. But which of these is an accurate description of the data may make a great difference in the theoretical conclusions an investigator would draw from the study.

The Frequency of Effects problem arises once it has been established that an experimental manipulation has produced an effect, and consists of trying to determine the frequency with which the effect occurred and the size of the effect when it occurred.

In order to illustrate one possible solution to this problem, we will, first, briefly describe an eye movement experiment in which the problem arises, second, describe a Frequency of Effects Analysis procedure, third, present the results of using the analysis in the study described, and finally, comment on some of the problems and uses of the analysis.

The Study: Variability in the Perceptual Span

There are a number of studies which have used eye movement contingent display control techniques to study the size of the visual region within which certain aspects of the text are perceived during a fixation in reading (McConkie & Rayner, 1976; Rayner, Inhoff, Morrison, Slowiaczek, & Bertera, 1981; Rayner. Well, & Pollatsek, 1980; Rayner, Well, Pollatsek, & Bertera, 1982). In these studies, subjects read from text displayed on a cathode ray tube (CRT) under computer control as their eye position was monitored. On selected fixations or on all fixations, depending upon the study, erroneous letters or masking stimuli occurred in selected areas of the text, with these areas defined with respect to the letter that was directly fixated during the fixation. The eye movement patterns were then examined to determine whether this change in the stimulus pattern in a particular visual

region caused a disruption in reading. If it did not, it was assumed that the type of information manipulated must not normally be used from that visual region during reading. These studies have agreed in finding that letter distinctions are perceived within a relatively small area, perhaps four character positions or less to the left of the fixated letter, and eight or fewer to the right. with lengths of words being perceived somewhat farther to the right.

The question studied in the experiment to be described here was this: Are these regions perceived in their entirety on each fixation, or does the actual region perceived vary from fixation to fixation, though always being within the perceptual spans observed in the earlier studies? The strategy used was similar to that just described: During selected fixations the letters at certain locations, defined with respect to the location of the fixated letter, were replaced by other letters, thus resulting in erroneous letters being present at specific retinal locations on those fixations. The eye movement data were then analyzed to determine whether the errors produced an effect, and if they did, to estimate the frequency with which this occurred.

The experimental manipulation was produced in the following manner. Text was displayed on the CRT, refreshed every 3 msec. The subject's eyes were monitored during reading, with sampling of eye position every msec. During the reading of each line, two fixations were selected as critical fixations. As soon as the location of the eyes on such a fixation could be determined, the display was changed in the manner

proscribed for that fixation.

Three of the conditions used will be described here, one control and two experimental. In the control condition, no change was made in the text, thus no errors were present during those selected fixations. In the Left-0 condition, all letters to the left of the fixated letter were replaced by other letters, with each letter being replaced by the letter least visually similar to it which did not change the external shape of the word. In the Right-3 condition, all letters more than 3 to the left of the fixated letter were replaced by other letters in the same manner. These erroneous letters remained in the text until the following eye movement began, at which time the normal text was returned to the screen. Thus, the erroneous letters were present in the text only during occasional single fixations. In the experimental conditions, these erroneous letters appeared at locations previously shown to be within the region in which letters are perceived during fixations in reading (Underwood & McConkie, 1985). The question to be investigated was whether these erroneous letters would produce an effect on reading every time they occurred or only on some occasions. Were the retinal areas where these letters appeared being attended during all fixations, or only during some of them?

Twelve college students served as subjects in the study, all having good reading skills. Each subject read 16 short passages of about 200 words each. This yielded over 1000 fixations in each condition on which the erroneous letters were present. The subjects were aware of the

occasional occurrence of the errors, but reported that they were not bothered much by them. In fact, some of the subjects went a quarter or half of the way through the experiment before becoming aware of the errors.

The dependent variables to be discussed here were the duration of the fixation on which the errors were present (fixation F0), the duration of the following fixation (fixation F1), and the length and direction of the intervening eye movement (saccade S1).

Initial analyses indicated that all three variables showed significant effects of the manipulations. Condition Left-0 increased the duration of fixation F0 by 21 msec, increased the duration of fixation F1 by 13 msec, and increased the likelihood of a regression on saccade S1 from 17% to 34%. Condition Right-3 had no effect on the duration of fixation F0 but increased the duration of fixation F1 by 21 msec and reduced the length of S1 forward saccades by .6 character position. While these differences indicated that effects were occurring, they did not indicate their frequency. This required an additional analysis.

Frequency of Effects Analysis

It is assumed that. had it not been for the experimental manipulation, the frequency distribution of the data from the experimental condition would be very similar to that from the control condition. The difference between the two distributions is assumed to

be due to the fact that on a certain proportion _p_ of the occurrences of the experimental manipulation an effect was produced on the dependent variable. The size of this effect is assumed to be normally distributed, with a mean of _e_ and a standard deviation of _v_. To simplify the model, the values of these parameters are treated as constant across subjects and across the equency distribution intervals.

Estimates of these parameters can be obtained through an iterative procedure in which different possible values are tried and their effects observed. This procedure involves selecting a value for each parameter, then modifying the frequency distribution from the control condition according to these parameters; that is, taking _p_ of the instances for each score or interval and increasing their value such that these instances are now normally distributed with a mean that is _e_ greater than their original value and a standard deviation of _v_. This is done for each score or interval. When completed, a new frequency distribution is created by grouping the resulting instances, including both those that were not modified and those that were. This new distribution is then compared to the actual frequency distribution obtained from the experimental condition by obtaining a sum of squared differences (SS_d) index. The combination of values for the three parameters which results in the lowest SS_d index is taken to be the best estimate of these parameters.

For our analyses we used the following procedures.

1. Decide on intervals to use in the frequency distribution. (We divided the range of fixation duration values into 20 msec intervals. with larger intervals at the extremes.) Obtain the frequencies for each interval for both distributions, (i.e., that from the experimental condition and that from the control condition).

2. Select a value for each of the parameters, _p_, _e_ and _v_. Values chosen on different iterations can either be selected for the purpose of sampling a wide range of combinations, or obtained by hill-climbing techniques to more efficiently seek the optimal combination.

3. Begin the creation of a third, hypothesized distribution by letting the frequency within each interval be equal to (1 - _p_) times the frequency of the corresponding interval in the control group distribution. This indicates the number of instances in each interval that are hypothesized to have remained unaffected by the experimental manipulation.

4. A certain proportion _p_ of the instances in each interval are assumed to have been affected by the manipulation. Thus, the scores for these instances must be increased as hypothesized. This was done by taking _p_ times the number of instances receiving each score in the control condition, and then distributing these instances as a normal curve with a mean that was _e_ msec greater than the original score with a standard deviation of _v_. The number

of these instances that now fell within each of the intervals was calculated, and was added to the intervals in the new distribution being created. This was done for each score present in the data, creating a new hypothesized distribution.

5. The hypothesized distribution was compared to the distribution actually obtained from the experimental condition by using the SS_d index. The difference was obtained between the frequencies of the hypothesized and experimental distributions for each interval, and these values were squared and summed. The resulting SS_d index was taken as an index of the degree of similarity between the two distributions.

6. This process was repeated with other combinations of values for the parameters until the combination which yielded the lowest SS_d index value was obtained. These values were taken as the best estimate of the frequency with which the experimental manipulation produced an effect, and the nature of that effect when it occurred (i.e., its average size and variance).

We found that the simplest approach was to begin by assuming that $v = 0$, that is, that the effect size does not vary. This makes the calculations much more rapid. Once the optimal values are found for the other two parameters, it is possible to explore the effects of manipulating v. In our analyses, adding the third parameter to the model made little difference to the optimal values obtained for p and e, though it did lead to a better fit as indicated by a reduced SS_d value.

Application of the Frequency of Effects Analysis to the Current Data

The Frequency of Effects Analysis was carried out with the F0 fixation duration data, where the experimental manipulation had been found to produce an average increase in duration of 21 msec. The frequency distributions for these fixation durations are shown in Figure 1. The best fit to the data was obtained with values of .21 for *P*, 151 for *e* and 31 for *v*. Thus. it is estimated that the experimental manipulation actually influenced the F0 fixation duration in 21% of the cases, while the size of the effect, when it occurred, averaged 151 msec with a standard deviation of 31 msec. By this estimate, in 79% of the cases no effect was produced in the Left-0 Condition.

Insert Figure 1 about here

Some confirmation of the accuracy of this estimate of the frequency of the effect was obtained when it was found that an effect occurred only for those fixations which were followed by a regressive eye movement. The mean duration of fixations which were followed by a forward eye movement was within 5 msec of the corresponding mean for the control condition. Furthermore, there was a 20% reduction in the number of forward saccades in the experimental condition compared to the control. If 20% of the eye movements in the experimental condition were regressions induced by the presence of erroneous letters, and it was

these instances which resulted in longer fixations, then these data suggest a pattern similar to the estimate obtained through the Frequency of Effects Analysis.

Figure 2 presents additional detail concerning the results of the Frequency of Effects Analysis showing the SS_d index for different values of the parameters _P_ and _e_. The graph on the left plots the sum obtained for different values of _e_, the effect size parameter, given the optimal value of _P_ for each value. The graph on the right is a similar plot, showing SS_d for different values of _P_, the frequency of effect parameter. when _e_ was optimized for each. It is clear that there is a certain region within which each parameter minimizes the SS_d. The effect of varying the third parameter, _v_. was to reduce the sum of squared deviations, but it had little effect on the shapes of the curves shown in Figure 2.

Insert Figure 2 about here

Using the data from durations of F1 fixations following forward saccades, a second application of the Frequency of Effects Analysis involved the data from the Right-3 condition. In this case, no effects were found on fixation F0, but a significant difference of 21 msec was found on fixation F1. The analysis yielded estimates of .36 for _P_, 54 for _e_, and 30 for _v_. Thus, it was concluded that, in the cases where

the F0 fixation showed no effect of the errors in the text, 36% of the F1 fixations showed an effect of an average 54 msec increase in the duration. with a 30 msec standard deviation.

Figure 3 presents the graphs of the sum of squared deviations as the parameters *p* and *e* were varied. Again, the graphs show clear regions where the index was minimized.

Insert Figure 3 about here

In the study described, other effects of the experimental manipulations were also identified. leading to the conclusion that readers seldom, if ever, fail to perceive letters in the regions manipulated.

A third application of the Frequency of Effects Analysis was carried out with the S1 forward saccade length data from condition Right-3. In this condition, the duration of fixation F0 showed no effect, but saccade S1 was shortened by an average of .6 character position. This analysis yielded an estimate of 1.0 for *p*, of .7 for *e* and .2 for *v*. These parameters suggest that all saccades were shortened by about .7 character positions. Thus, again there is no indication of some instances in which erroneous letters in this region went undetected.

The application of the Frequency of Effects Analysis to these data appeared successful. It yielded quite different results for different sets of data, and the plots of the sum of squared deviations showed clear areas where the SS_d index was minimized. The results suggested quite a different picture of the effects produced by the experimental manipulation than might have been assumed without its use. Finally, in one case, independent data confirmed the frequency of effect estimate given by the analysis.

Comments on the Frequency of Effects Analysis

Knowledge of the characteristics of the Frequency of Effects Analysis is limited at the present time. Monte Carlo investigation is required in order to estimate the sampling characteristics of the parameters, to determine the sample size needed to yield stable parameter estimates, and to establish how robust the Frequenct of Effects Analysis is to violations of its assumptions. In the meantime, the analysis should probably be used only with large data sets (in the data reported, sample sizes ranged from 700 to over 1000 for each condition) and the assumptions should be noted carefully.

It should be recognized that the particular model used in this analysis is only one of many possible models. It may be necessary to modify the model in certain cases to bring it in harmony with known or suspected characteristics of the data.

In spite of the limitations and concerns that exist at the present

time, the general technique appears to be useful in revealing characteristics of experimental data which are not apparent from techniques typically used. There are many circumstances in which it would be useful to be able to estimate the frequency of effects in eye movement research. This analysis would be appropriate for use with various eye movement measures, including fixation durations, saccade lengths, and viewing time indices such as gaze duration and reading time. In addition, it would sometimes be appropriate to use the Frequency of Effects Analysis in studies where the time to accomplish certain tasks, such as finding a target in a complex display, serves as the dependent variable. Finally, it should be useful whenever a researcher needs to know the frequency with which an effect occurred.

References

McConkie, G. W., & Rayner, K. (1976). Asymmetry of the perceptual plan in reading. *Bulletin of the Psychonomic Society*, *8*, 365-368.

Rayner, K., Inhoff, A. W., Morrison, R. E., Slowiaczek, M. L., & Bertera, J. H. (1981). Masking of foveal and parafoveal vision during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 167-179.

Rayner, K., Well, A. D., & Pollatsek, A. (1980). Asymmetry of the effective visual field in reading. *Perception & Psychophysics*, *27*, 537-544.

Rayner, K., Well, A. D., Pollatsek, A., & Bertera, J. H. (1982). The availability of useful information to the right of fixation in reading. *Perception & Psychophysics*, *31*, 537-550.

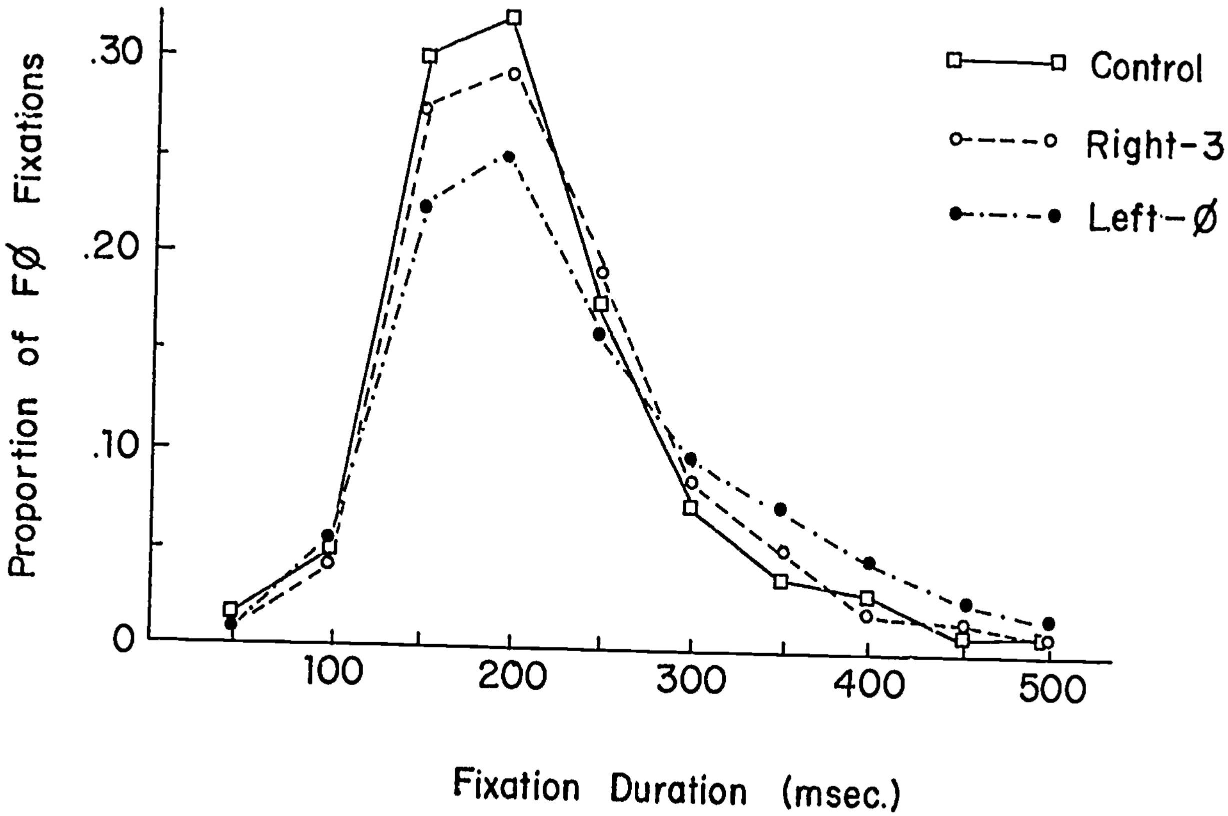
Underwood, N. R., & McConkie, G. W. (1985). Perceptual span for letter distinctions during reading. *Reading Research Quarterly*. *20*, 153-162.

Figure Captions

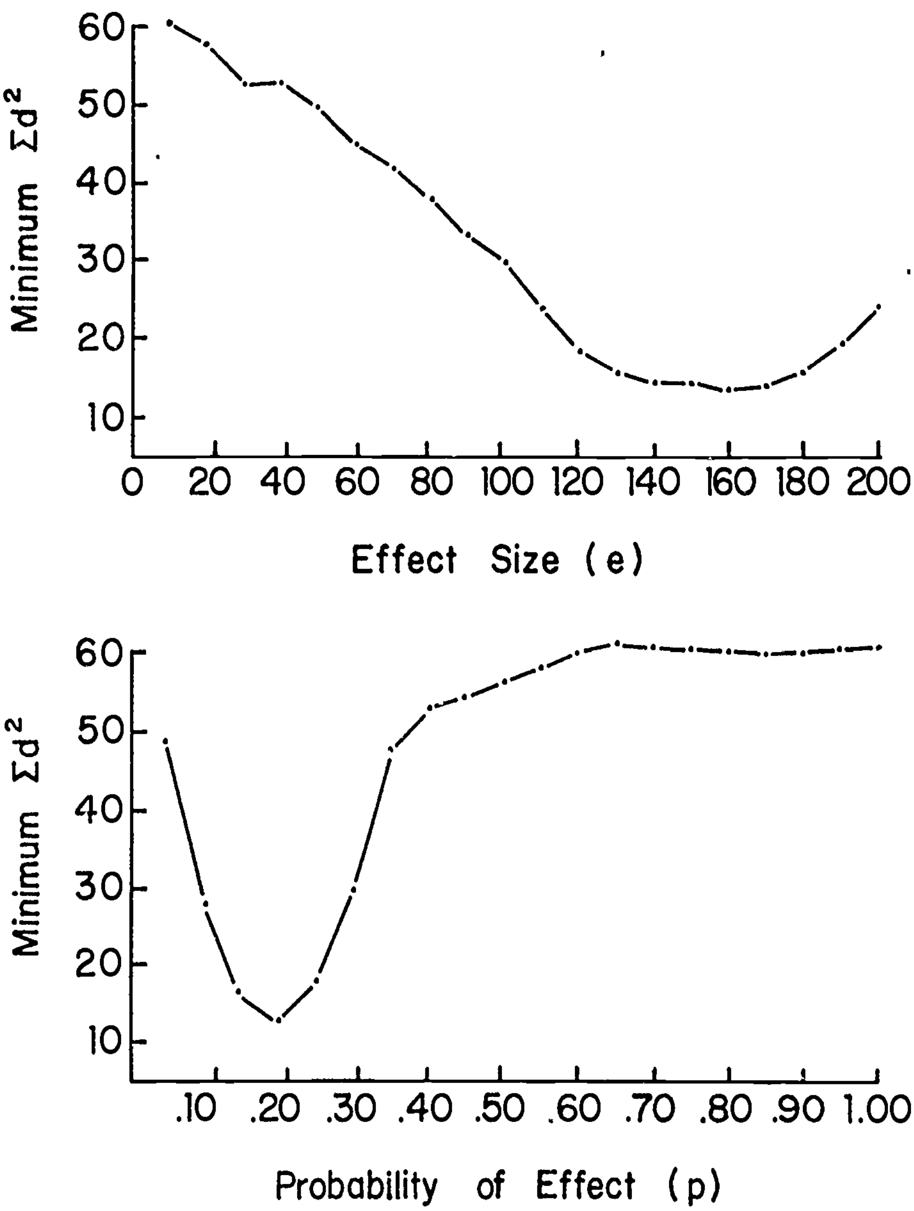
Figure 1. Frequency Distributions for the Duration of Fixation F0, the Critical Fixation, in Control, Left-0 and Right-3 Conditions.

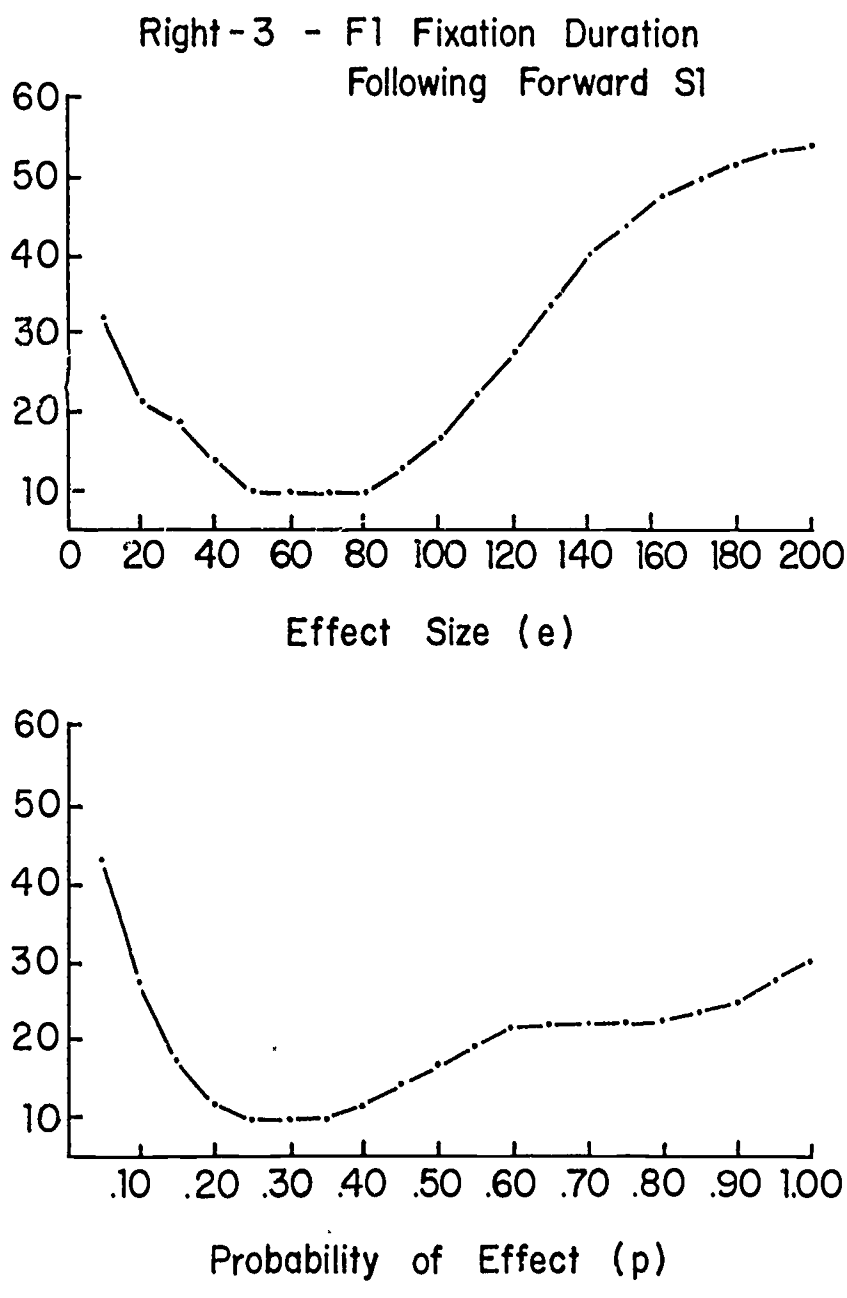
Figure 2. SS_d Values from Frequency of Effects Analysis of Fixation F0 Duration for Condition Left-0.

Figure 3. SS_d Values from Frequency of Effects Analysis of Fixation F1 Durations for Condition Right-3.


Proportion of FØ Fixations
.30
.20
.10
0
100
200
300
400
500
Fixation Duration (msec.)
Control
Right-3
Left-Ø

Left-O - F.O Fixation Duration




Right-3 - F1 Fixation Duration
Following Forward S1
60
50
40
30
20
10
0 20 40 60 80 100 120 140 160 180 200
Effect Size (e)
60
50
40
30
20
10
.10 .20 .30 .40 .50 .60 .70 .80 .90 1.00
Probability of Effect (p)